



Contextual Recommendation of Social Updates, a Tag-based Framework

Adrien Joly, Pierre Maret, Johann Daigremont

► To cite this version:

Adrien Joly, Pierre Maret, Johann Daigremont. Contextual Recommendation of Social Updates, a Tag-based Framework. Active Media Technology 2010, Aug 2010, Toronto, Canada. pp.436-447. hal-00531366

HAL Id: hal-00531366

<https://hal.science/hal-00531366>

Submitted on 3 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Contextual Recommendation of Social Updates, a Tag-based Framework

Adrien Joly^{1,2}, Pierre Maret³, and Johann Daigremont¹

¹ Alcatel-Lucent Bell Labs France,
Site de Villarceaux, F-91620 Nozay, France
`first.last@alcatel-lucent.com`

² Université de Lyon, CNRS,
INSA-Lyon, LIRIS, UMR5205, F-69621, France

³ Université de Lyon, Laboratoire Hubert Curien, UMR CNRS 5516,
F-42000 Saint-Etienne, France
`first.last@univ-st-etienne.fr`

Abstract. In this paper, we propose a framework to improve the relevance of *awareness* information about people and subjects, by adapting recommendation techniques to *real-time web* data, in order to reduce *information overload*. The novelty of our approach relies on the use of contextual information about people’s current activities to rank *social updates* which they are following on Social Networking Services and other collaborative software. The two hypothesis that we are supporting in this paper are: (i) a social update shared by person X is relevant to another person Y if the current context of Y is similar to X ’s context at time of sharing; and (ii) in a web-browsing session, a reliable current context of a user can be processed using metadata of web documents accessed by the user. We discuss the validity of these hypothesis by analyzing their results on experimental data.

1 Introduction

On Social Networking Services (such as Facebook⁴, Twitter⁵, LinkedIn⁶) and other collaboration software, people maintain and create new social ties by sharing personal (but not necessarily private) *social updates* regularly to their community, including *status messages* and *bookmark notifications*. As depicted on Figure 1, a *social update* is a short message sent to a group of interested persons (e.g. a community). It can consist of a one-sentence news or question, an anchor, a picture, or a comment, to share their current thoughts, activities, intentions and needs.

On most of these tools, social updates are not meant to be consumed in a *push* fashion, i.e. like emails, that are aimed at specific recipients and stacked in their inboxes. Instead, community members can go through the list of short social

⁴ <http://www.facebook.com/>

⁵ <http://www.twitter.com/>

⁶ <http://www.linkedin.com/>

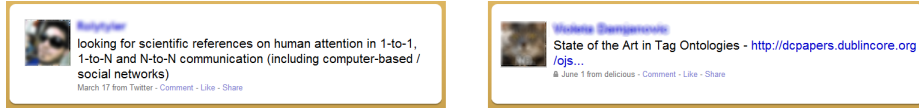


Fig. 1. A status update from Twitter, and a bookmark update from Delicious.

updates of the people or subjects (e.g. *hashtags* on twitter) they follow, to get a quick feeling of *awareness* about those they care about. However, as the number of people and subjects being followed increases, the time required to get through to the social updates they emit also increases, causing a loss of productivity. Additionally, as social updates are broadcast in real-time, they create frequent interruptions that can reduce people’s ability to focus on a demanding task, especially when the social updates are not relevant for this task (because it would induce a costly cognitive switch).

In response to this emerging problem of *information overload*, we propose a framework to rank *social updates* according to real-time distances in-between users’ contexts, which are processed on social and meta-descriptions of the web documents the users are looking at. The two underlying hypothesis that we are supporting in this paper are: (i) a social update shared by person X is relevant to another person Y if the Y ’s current context is similar to X ’s context at time of sharing; and (ii) in a web-browsing session, a reliable current context of a user can be processed using a combination of tags and metadata of web documents accessed by the user.

In the next section, we motivate our approach by explaining how context-awareness and current web techniques can be leveraged to improve awareness. In section 3, we survey existing work related to our problem. In section 4, we describe our contextual recommendation framework to provide relevant social updates to people. In section 5, we present our experimental setup and results to evaluate the human response to this approach. We will then discuss these results and propose future work.

2 Motivation

Vyas et al. proved [18] that informal ties between co-workers are essential to improve awareness, and thus better collaboration. In previous studies [9, 8], we have identified that *contextual information* about users could be leveraged to assist the sharing of social updates, and thus maintain these ties, while reducing interruptions. *Context* was defined by Dey [5] as “*information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves*”. In most context-aware applications, researchers have been relying on sensors to extract contextual information. As tagging becomes a common practice on the Internet, rich contextual information can also emerge from human-generated content.

Despite some semantic issues related to ambiguity of terms combined in folksonomies, the increasing amount of tags given by Internet users on digital resources [12, 11] (e.g. web pages tagged on delicious⁷) have become good indexing features for recommender systems [7, 15]. With the growing use of twitter and geotagging applications in mobility, tags are now emerging from places, events and other real-world entities [13], which implies exciting opportunities to create new ambient intelligence, ambient awareness, augmented reality, and other social applications.

3 Background

To the best of our knowledge, the closest existing solution to our problem is a web-based service and mobile application called *My6sense*⁸. This software can filter entries from RSS and other feeds (including social updates) according to the user's preference. This content-based filtering technique relies on a profile which contains the user's subjects of interest, and this profile is continuously evolving by tracking which entries are consulted by the user. Similarly, *SoMe-ONe* [1] is a collaboration tool that can recommend contacts (and their selection of web resources) to users by identifying communities of interest around topics, using collaborative filtering. The names of users' bookmark categories are leveraged as topics. *groop.us* [3] applied the same approach while relying on tags attributed to web pages by users (folksonomies from a social bookmarking website) instead of hierarchical categories. In these approaches, recommendations are based on documents that were explicitly selected and shared (bookmarked) by users. Despite the evolving design of user profiles, the filtering is not adaptive to the user's current context.

It is possible to provide collaboration opportunities by recommending people that are currently browsing similar documents [4, 2], based on a TF-IDF analysis [17] of their content, users' context being represented by weighted term vectors. These recommendations can also include some information about people's current activities, as identified by a software module that tracks users' actions on their computer (e.g. chat sessions, office documents being edited, etc...) [6]. However these efforts don't leverage tags proposed by web users.

In the PRUNE framework [10], contextual entities (e.g. person, place or resource) and events can be extracted from heterogeneous sources like RSS feeds, web APIs and manual user entry. The *Notes that Float* application leverages this framework to attach such contextual information to notes added by the user, so that their individual visibility depends on their relevance with the user's current content, which relies on their similarity with the context at the time these notes were added. However, we have found no evidence that tags were leveraged in this application. Moreover, previous collaborative systems imply potential privacy issues.

⁷ <http://del.icio.us/>

⁸ <http://www.my6sense.com/>

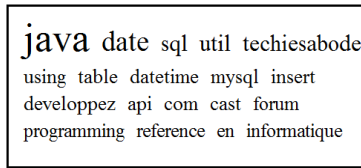


Fig. 2. Sample contextual tag cloud: a set of words describing one's context

4 Contextual recommendation framework

After having identified that contextual information about people can be leveraged to further describe the documents they are browsing/editing, and thus to recommend these documents to people that are in a similar context, we have reviewed several techniques and applications that are relevant for ranking documents.

In this section, we study the case of enterprise employees working on computers, then we present a framework and software implementation of an adapted social update recommender system which considers web-browsing context as a relevance criteria, and leveraging their tags as features.

4.1 Case study

As a motivating case, we propose to consider an enterprise environment where employees work on individual networked computers. They don't know everybody in the enterprise, and are possibly spread across several offices in several cities, or even different countries. Such organizations traditionally rely on hierarchies of managers to coordinate the efforts of workers and transfer information to the relevant parties.

We propose an internal social networking tool that allows every worker to share and retrieve relevant information about the current interests and status of their colleagues, while reducing unnecessary task interruptions and network maintenance time. This system will rely on various stream/feed to leverage personal current thoughts, activities, intentions and questions, and must respect their privacy (e.g. case of private browsing).

4.2 Contextual tag clouds

As users are working on computers in the case study presented above, most contextual information about their current activity can be extracted from the software they use (e.g document edition and viewing). In this study, we assume that descriptions about the web sites they are currently browsing (e.g. to find some reference on an ongoing task), can provide clues on the user's current activity. Users' context can thus be modeled as a set of weighted terms, combining metadata and tags about these browsed web pages. As these terms can potentially reveal private or confidential information, users must be able to quickly

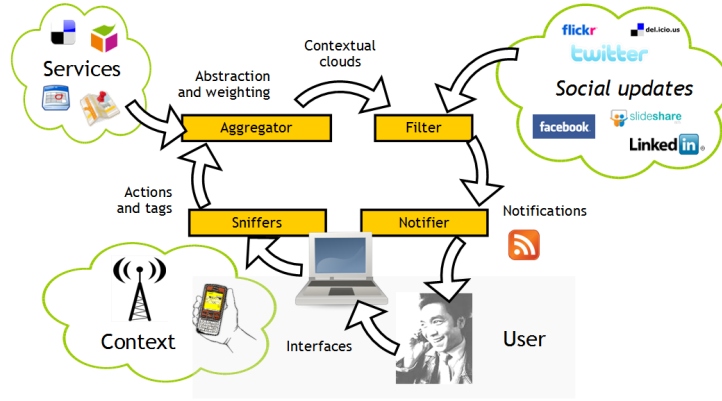


Fig. 3. Overview of the contextual recommendation loop

visualize and edit them before submitting to a recommender system. We propose the name of "*Contextual Tag Clouds*" to refer to these human-readable contexts based on a weighted combination of tags and other descriptive terms, as depicted on Figure 2.

4.3 Data flow and user interaction

As depicted on Figure 3, contextual information is extracted from user-manipulated content (in our case, descriptions of web pages currently browsed) by *sniffers* running on every user's computer. For privacy control reasons, no contextual information will be ever sent to any remote party without user confirmation. Contextual information is represented by a set of weighted keywords, and represented as a *tag cloud*.

Also running on the user's computer, an *aggregator* gathers events from all these sniffers, and queries several web services to generate weighted tags, in order to combine them in a *contextual tag cloud* that represents the user's current context.

When posting a social update (e.g. a *tweet*), the user can attach his/her current contextual tag cloud, so that the *contextual filter* (i.e. a recommender system running in the infrastructure) can match it with other users' last contextual tag cloud, using a relevance function.

This social update will then be recommended to users whose last contextual tag cloud is relevant to the one attached to the social update. That way, every user gets a dynamic (near real-time) list of recent social updates, sorted by decreasing relevance score as they browse the web. Like with regular social networking and microblogging services, these short social updates can be quickly read by users to remain *aware* of relevant activities going on in their communities. They can also decide to reply to social updates or to call their author.

4.4 Ranking model

The theoretical framework that we designed to solve our relevance ranking problem relies on a vector space model, five weighting functions, an aggregation operator and a relevance function. The weighting functions are equivalent to *context interpreters* [5]: they transform raw data (in our case, URLs of web documents being browsed by the user) into higher-level information (contextual tags) by abstraction.

The *contextual tag cloud model* is equivalent to the vector space model proposed by Salton [17]. Notice that, in this paper, the word *tag* refers to terms, whatever their origin. Traditionally, a set of terms resulting from the analysis of a document d is formalized as a vector of weights $v(d) = [w(t_1, d), w(t_2, d), \dots, w(t_N, d)]$ attributed for each term $t = [t_1, t_2, \dots, t_N \in \mathbb{R}]$. The specificity of our model lies in the combination of five functions applied on browsed documents (and their crowd-sourced description: *tags*) to compute the weights:

1) the *Metadata* function counts occurrences of term t with different coefficients (α, β, γ) , depending of the position of this term in document d 's metadata:

$$w_1(t, d) = \alpha * |t \in T_d| + \beta * |t \in K_d| + \gamma * |t \in D_d|$$

where $|t \in T_d|$ is the number of occurrences of the term t in the *title* of the document d , $|t \in K_d|$ in its *keywords* set, and $|t \in D_d|$ in its *description* text.

2) the *SearchQuery* function counts the number of occurrences of term t in a search query Q_d , when the analyzed document d contains a list of search results:

$$w_2(t, d) = |t \in Q_d|$$

3) the *DomainNames* function adds the domain names (including sub-domains) N_d from document d 's URL as terms:

$$w_3(t, d) = |t \in N_d|$$

4) the *SocialBookmarks* function counts the number of people who publicly bookmarked the document d using the term t as tag:

$$w_4(t, d) = \sum_{p \in P} \text{tag}(p, d, t)$$

where each person p is in the set of people P that are using this bookmarking service, and where $\text{tag}(p, d, t)$ has a value of 1 or 0, whether or not this person p bookmarked the document d using term t as a tag.

5) the ***SemanticAnalyzer*** function counts the number of occurrences of semantically-defined entities (i.e. concepts and instances) that are represented by the term t , when they are identified in the document d :

$$w_5(t, d) = |t \in R_d|$$

$$R_d = [\forall e \in E_d, repr(e)]$$

where $repr(e)$ is the textual representation of a semantic entity e , R_d is the set of textually represented entities E_d found in the document d . This function is further described in the next part of this section.

Additionally, we define an aggregation operator and a relevance function that leverage the vectors resulting from the weighting functions above:

The ***aggregation operator*** is the addition of given weighted term vectors, after their individual normalization. The normalized form $\|v\|$ of vector v conforms to $\sum_{t=t_1}^N \|v_t\| = 1$, with weight values $v_t \in \mathbb{R}$ in the range $[0, 1]$. Thus, the aggregation operator applied to a set of vectors $V = [v_1, v_2, \dots, v_M]$ acts as the following function:

$$aggr(V) = \left\| \sum_{t=1}^M \|v_t\| \right\|$$

The ***relevance function*** between normalized vectors, like in traditional vector-based models, relies on cosine similarity. Thus, the relevance of a tag cloud vector R with another tag cloud vector S is computed by:

$$relevance(R, S) = \frac{R \cdot S}{\|R\| \|S\|}$$

which returns a relevance score $r \in \mathbb{R}$ in the range $[0, 1]$, 1 being the maximum relevance score (i.e. contextual equality).

4.5 Software implementation

The framework described above was designed as a modular architecture, according to the data flow depicted in Figure 3, in which software modules communicate through RESTful HTTP requests. In this section, we present the implementation of these modules:

- A Firefox extension⁹ acts as a *context sniffer* and a *notifier*. For sniffing, it hooks on the browser's events related to opening, closing and switching web pages, and transmits these events with the corresponding URLs to the local *Context Aggregator* for processing. At the end of the flow, the recommended social updates are displayed in the side-bar of the browser.

⁹ <http://www.firefox.com/>

- The *Context Aggregator* handles local events with their attached contextual information, and runs weighting functions on this information to produce an aggregated (and thus normalized) contextual tag cloud for the *Contextual Filter*, using the *aggr()* function defined in the previous section. The *weighting functions* are implemented as five interpreters that turn URLs into contextual clouds (i.e. weighted term vectors). The *Metadata* interpreter parses the *title*, *description* and *keywords* elements from the HTML source code of each web page to produce the corresponding weighted terms, with the following parameter values: $\alpha = 50$ per term appearing in the title, $\beta = 10$ in the *keywords* field, and $\gamma = 1$ in *description* field. The *SearchQuery* interpreter extracts query terms from Google Search¹⁰ result pages. The *Social-Bookmarks* interpreter gathers tags given by users about a web page, when existing on the Delicious social bookmarking service. The *SemanticAnalyzer* gathers textual representations of semantic entities that were identified in the web page, thanks to the SemanticProxy web service¹¹.
- The *Contextual Filter* receives contextual clouds gathered and interpreted by users' *aggregator*, computes relevance scores between them using the *relevance()* function, and recommend best-ranked social updates to each user (through their *notifier*). Social updates are gathered by subscription to the users' declared third-party social feeds/streams (e.g. their Twitter account).

This software ecosystem is functional and gives a good sense of the benefits of our approach. In the next section, we present an evaluation of the underlying framework.

5 Evaluation

In order to evaluate the validity of our hypothesis on relevance of contextually recommended social updates, we gathered browsing logs and social updates from 8 volunteers during one week, ran our algorithms on these logs to generate 1846 contextual clouds (every 10 minutes), and asked the volunteers to rank the quality of a selection of social updates. In this section, we define the experimentation plan we followed, explain its setup, then discuss the results obtained.

5.1 Experimentation plan

The evaluation of our hypothesis relies on two measures: (i) the relevance of social updates with the context of their author at the time of posting, and (ii) their relevance for other users in similar contexts. As the quality of recommendations is to be evaluated by users with their own browsing behavior, implied contexts, and own social updates, we did not rely on existing evaluation data sets such as the ones from TREC, nor follow a scenario-based experiment.

¹⁰ <http://www.google.com/>

¹¹ <http://semanticproxy.opencalais.com/>

During one week, volunteers browsed web pages using Firefox and produced social updates (i.e. shared statuses and bookmarks), while the provided sniffing extension was logging the required browsing events to a local database. At the end of this period, they were proposed to review these log entries, so that they could remove privacy-critical entries when needed (e.g. private activities, and other noisy data that is irrelevant to this study), and then send us their log.

Afterwards, we ran our algorithms on the browsing logs and social updates provided by volunteers, to produce personalized survey forms containing ranked recommendations for each volunteer. We asked each volunteer to fill two personalized surveys. In the first survey, we asked volunteers to rate the perceived relevance of three random social updates with five contextual clouds generated from their own web browsing data. For each context, only one of the proposed social updates was actually a well-ranked match. In order to support them in remembering those contexts, we provided the list of web pages that were being browsed by the volunteer at that time. In the second survey, volunteers rated the relevance of their own social updates with their contextual tag cloud at the time of posting.

5.2 Experimental setup and process

In this section we provide the process and parameters that we set to generate these personalized surveys from the logs provided by volunteers.

Because the experiment was not interactive, we indexed contextual clouds and social updates on a common time line with a period of 10 minutes. Contextual clouds are generated from the list of URLs involved in a web browsing event, i.e. when the page was opened, selected or closed. Indexing a social update consists of associating it with the contextual cloud of the last context snapshot at the time of posting this update. If there is no known context information in the previous snapshot, we use the one before the previous. Every indexed contextual cloud is processed to split multiple-word tags, cleaned from punctuation and other non-literal characters, filtered against a stop-words list, and then normalized so that the sum its tags' weights equals 1. Only the first 20 tags (with highest weights) are displayed to volunteers. As shown on Figure 2, a contextual tag cloud can contain diverse kinds of terms, such as words in various languages, word combinations and acronyms.

Then, we ran the recommendation algorithm on the contextual and social indexes in order to produce a relevance matrix for each participant. In order to generate a participant's personalized survey, we selected 5 heterogeneous contexts (i.e. the most dissimilar to each other) that were matched (by the recommender) with at least one highly-ranked social update. The second survey was simply generated by correlating users' social updates with their corresponding context.

5.3 Results

As stated above, the results are twofold: we gathered scores given by every participant on (i) the relevance of social updates with the context of their posting, and (ii) the relevance of social updates for other people with similar contexts.

Relevance of contextualized social updates: In order to measure the consistency of contextual clouds as reference documents for recommending social updates, we asked the participants to rate the relevance of each of their own social updates (e.g. their tweets and bookmarks) to the contextual cloud representing their current situation at time of posting/sharing.

Over a total of 59 social updates, their authors rated an average relevance to context of 50.3%. The following distribution of ratings is observed: 19 social updates were ranked 1 (low relevance), 10 were ranked 2, 14 were ranked 3, and 16 were ranked 4 (high relevance). These social updates are gathered from several social streams: 54% are status updates posted on Twitter, 29% are bookmarks instantly shared through Delicious.

By further analyzing these specific types of social streams, we discovered an average relevance score of 71% for shared bookmarks, and 38% for status updates from Twitter. It is natural that new bookmarks are more relevant to their context, as the web document that is bookmarked is usually being browsed by the user, and thus represented in the corresponding contextual clouds. Concerning status updates, Naaman et al. [14] proved that only 41% of social updates from twitter are actual statuses about the current activity of the person (categorized as "me now" by the authors). The similarity of this proportion with our average contextual relevance score for status updates gives some proof, although preliminary, about the consistency of our results.

Relevance of recommendations: As explained in the previous section, social updates proposed to users are voluntarily not all relevant. Our goal is to observe a correlation between the relevance scores given by participants and the rankings computed by the system. Thus, we rely on a Mean Percentage Error (based on MAE, Mean Absolute Error) to define the following *accuracy* function:

$$accuracy = 1 - \sum_{q=1}^Q |relevance(C_q, U_q) - rating(C_q, U_q)|$$

in which, for each proposed social update q , $relevance(C_q, U_q)$ is the relevance score of the social update U_q with the contextual tag cloud C_q , as evaluated by the ranking algorithm. Whereas, $rating(C_q, U_q)$ is the actual relevance score, as given by the volunteer. Both scores are values in the range $[0, 1]$, represented as percents. As $rating()$ scores are given by volunteers in the $[1, 4]$ grade rating range, they are converted to percents with the following formula:

$$rating = \frac{(grade - 1)}{3}$$

We observed an average accuracy value of 72%. As a natural behavior of recommender systems, the best-ranked ratings (mostly in *Rank 3*) are slightly overestimated by the recommendation algorithm, whereas low relevance ratings (*Rank 1*) given by participants are higher than expected.

From the list of relevance ratings expected by the recommendation system, 63% are low ranked (*Rank 1*), whereas 19% are medium-high (*Rank 3*). The high number of low-ranked scores and the medium ranking of better scores expected by the algorithm reveals that highly similar contextual clouds were rare in our small scaled experiment. By increasing the number of participants, more similar contexts would be found, thus the average scores would naturally increase.

6 Conclusion

In this paper, we proposed a theoretical framework, a privacy-aware implementation and its evaluation to rank social updates by contextual relevance, for reducing information overload. Through the analysis of experimental results, we evaluated a combined weighting scheme based on social and meta-descriptions of web pages being accessed by the user, as a contextual criteria for recommending relevant social updates. This study explores the potential of our novel recommendation approach based on contextual clouds. Despite the small scale of this preliminary experiment, our results are promising. The average accuracy of recommended social updates: 72%, is significant for a web recommender system. We observed that the relevance perceived by users increases as social updates reflect the current activity of their authors.

In order to improve the performance of our system, we intend:

- to improve the quality of context with emergent semantics of tags [16].
- to broaden the range of context, by developing additional context sniffers, including documents, and physical context information from mobile devices.
- to find more precise relevance factors between specific types of social updates and contextual properties, after having carried out a higher-scale experiment.
- and to improve the scalability of the system when used simultaneously by numerous users (currently: $O(n^2)$ complex), e.g. using tag clustering.

References

1. Agosto, L.: Optimisation d'un Réseau Social d'échange d'Information par Recommandation de Mise en Relation. Ph.D. thesis, Université de Savoie, France (2005)
2. Bauer, T., Leake, D.B.: Real time user context modeling for information retrieval agents. In: CIKM '01: Proceedings of the tenth international conference on Information and knowledge management. pp. 568–570. ACM, New York, NY, USA (2001)
3. Bielenberg, K., Zacher, M.: Groups in social software: Utilizing tagging to integrate individual contexts for social navigation. Digital Media. Bremen, Germany, University Bremen. Master of Science in Digital Media 120 (2005)

4. Budzik, J., Fu, X., Hammond, K.: Facilitating opportunistic communication by tracking the documents people use. In: Proc. of Int. Workshop on Awareness and the WWW. ACM Conference on CSCW 2000, Philadelphia. Citeseer (2000)
5. Dey, A.K.: Providing Architectural Support for Building Context-Aware Applications. Ph.D. thesis, Georgia Institute of Technology (2000)
6. Dragunov, A., Dietterich, T., Johnsrude, K., McLaughlin, M., Li, L., Herlocker, J.: TaskTracer: a desktop environment to support multi-tasking knowledge workers. In: Proceedings of the 10th international conference on Intelligent user interfaces. pp. 75–82. ACM New York, NY, USA (2005)
7. Hotho, A., Jschke, R., Schmitz, C., Stumme, G.: Information retrieval in folksonomies: Search and ranking. In: The Semantic Web: Research and Applications, pp. 411–426 (2006)
8. Joly, A.: Workspace Awareness without Overload: Contextual Filtering of Social Interactions. In: Smart Offices and Other Workspaces, workshop of the Intelligent Environments 2009 conference. pp. 297–304. Ambient Intelligence and Smart Environments, IOSPress (2009)
9. Joly, A., Maret, P., Daigremont, J.: Context-Awareness, the Missing Block of Social Networking . International Journal of Computer Science and Applications 4(2) (2009), special Issue on Networking Mobile Virtual Knowledge
10. Kleek, M.V., Karger, D.R., mc schraefel: Watching through the web: Building personal activity and Context-Aware interfaces using web activity streams. In: Proceedings of the Workshop on Understanding the User - Logging and Interpreting User Interactions in Information Search and Retrieval (UIIR-2009), in Conjunction with SIGIR-2009, Boston, MA, USA (2009)
11. Marlow, C., Naaman, M., Boyd, D., Davis, M.: Position paper, tagging, taxonomy, flickr, article, toread. Collaborative Web Tagging Workshop at WWW'06 pp. 31–40 (2006), <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.74.8883>
12. Mathes, A.: Folksonomies-cooperative classification and communication through shared metadata. Computer Mediated Communication (2004)
13. Naaman, M., Nair, R.: ZoneTag's collaborative tag suggestions: What is this person doing in my phone? Multimedia, IEEE 15(3), 34–40 (2008)
14. Naaman, M., Boase, J., Lai, C.: Is it really about me?: message content in social awareness streams. In: Proceedings of the 2010 ACM conference on Computer supported cooperative work. pp. 189–192. ACM, Savannah, Georgia, USA (2010)
15. Niwa, S., Doi, T., Honiden, S.: Web page recommender system based on folksonomy mining for itng'06 submissions. In: Information Technology: New Generations, 2006. ITNG 2006. Third International Conference on. pp. 388–393 (2006)
16. Rattenbury, T., Good, N., Naaman, M.: Towards automatic extraction of event and place semantics from flickr tags. In: SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. pp. 103–110. ACM, New York, NY, USA (2007)
17. Salton, G., McGill, M.J.: Introduction to Modern Information Retrieval. McGraw-Hill, Inc., New York, NY, USA (1986)
18. Vyas, D., Van De Watering, M., Elins, A., Van Der Veer, G.: Engineering Social Awareness in Work Environments. Lecture Notes in Computer Science 4555, 254 (2007)